Assignment Overview

Sunday, September 18, 2022   4:04 PM

In this module, begin working with the cleaned dataset from the previous module. You will now begin to analyze the dataset to find the distribution of data, presence of outliers and the correlation between different columns.

- ### Learning Objectives

  - Identify the distribution of data in the dataset.
  - Identify outliers in the dataset.
  - Remove outliers from the dataset.
  - Identify correlation between features in the dataset.

Assignment Overview

- After cleaning the dataset, your next step is the analysis. In this stage you will become more familiar with the data set and it will start to take shape. In this assignment you will:

  1. Plot a distribution curve, and histogram.
  2. Find the median, and outliers of particular columns.
  3. Compute the Inter Quartile Range.
  4. Find out the upper and lower bounds, and find correlations between numerical columns.
  5. Create a new dataframe.